

A high dimensional Central Limit Theorem for martingales, with applications to context tree models

Roberto Imbuzeiro Oliveira¹

¹ IMPA

We establish a central limit theorem for (a sequence of) multivariate martingales which dimension potentially grows with the length n of the martingale. A consequence of the results are Gaussian couplings and a multiplier bootstrap for the maximum of a multivariate martingale whose dimensionality d can be as large as e^{n^c} for some $c > 0$. We also develop new anti-concentration bounds for the maximum component of a high-dimensional Gaussian vector, which we believe is of independent interest. The results are applicable to a variety of settings. We fully develop its use to the estimation of context tree models (or variable length Markov chains) for discrete stationary time series. Specifically, we provide a bootstrap-based rule to tune several regularization parameters in a theoretically valid Lepski-type method. Such bootstrap-based approach accounts for the correlation structure and leads to potentially smaller penalty choices, which in turn improve the estimation of the transition probabilities.

Joint work with Alex Belloni (Duke)